Theoretical and practical aspects of linear optimization

Tijssen, G.A.

Chapter 4

B-T Matrices and Applications of B-T Tableaus

4.1 Introduction

In chapter 3 we introduced the concept of B-T tableau. In this chapter we will go into more detail and establish several more properties related to these tableaus. Furthermore, we will give a number of examples that show their usefulness.

4.2 B-T Matrices

B-T tableaus are special forms of Tucker Simplex Tableaus. In case of Simplex tableaus the objective row (the row corresponding to the objective function) and the right-hand-side(rhs) column have a special meaning, and no pivots are allowed on entries from it. Consider a B-T tableau of an LO-problem in which both the objective row and the rhs column are all-zero vectors. If we remove the objective row and the rhs column we obtain a tableau that we will call a Balinski-Tucker(B-T) matrix. Actually, it is a tableau that represents the constraint collection $Ax_N = -x_B, x_N \geq 0, x_B \geq 0$ with the rhs-column removed. As usual, x_B is the vector of basis variables, and x_N the vector of nonbasic variables.

A more general formulation reads as follows. A real matrix A with rows labeled $1, \ldots, m$ and columns labeled $1, \ldots, n$ is called a B-T matrix iff it has the following properties:

- A contains an all-zero submatrix with rows $1, \ldots, r$ and columns $1, \ldots, q$ with $0 \leq r \leq n$ and $0 \leq q \leq m$. The values of r and q are not both zero.
- The submatrix of A with columns $r+1, \ldots, n$ and rows $1, \ldots, q$ consists of lexicographically positive columns. The columns in this submatrix are lexicographically nondecreasingly ordered.
- The submatrix of A with rows $q+1, \ldots, m$ and columns $1, \ldots, r$ consists of lexicographically negative rows. The rows in this submatrix are lexicographically nonincreasingly ordered.

The general form of a B-T matrix is depicted in Figure 4.1. Note that if $r = 0$, q can be taken equal to m since all columns are lexicographically positive.

The central problem in this section is called the *B-T matrix problem* (BT-MAT). It is the problem of transforming an arbitrary matrix into a B-T matrix by means of the following three operations:

ER Exchange two columns c_1 and c_2; $1 \leq c_1 < c_2 \leq n$.

EC Exchange two rows r_1 and r_2; $1 \leq r_1 < r_2 \leq m$.

PIV 'Exchange' a row p with a column c by means of a pivot operation on a nonzero entry of A, say A_{pc}; $1 \leq p \leq m$ and $1 \leq c \leq n$.

A pivot operation on a nonzero entry A_{pc} transforms the matrix A into an equivalent matrix by means of the following rules:

$$\begin{array}{ll} A_{ij} := A_{ij} - (A_{ic} \times A_{pj})/A_{pc} & i = 1, \ldots, m, i \neq p \\ & j = 1, \ldots, n, j \neq c \\ A_{pj} := A_{pj}/A_{pc} & j = 1, \ldots, n, j \neq c \\ A_{ic} := -A_{ic}/A_{pc} & i = 1, \ldots, m, j \neq p \\ A_{pc} := 1/A_{pc} & \end{array}$$

Note that this is nothing else than a pivot operation used in the Simplex method on a Tucker tableau with '$\leq$' constraints. In terms of

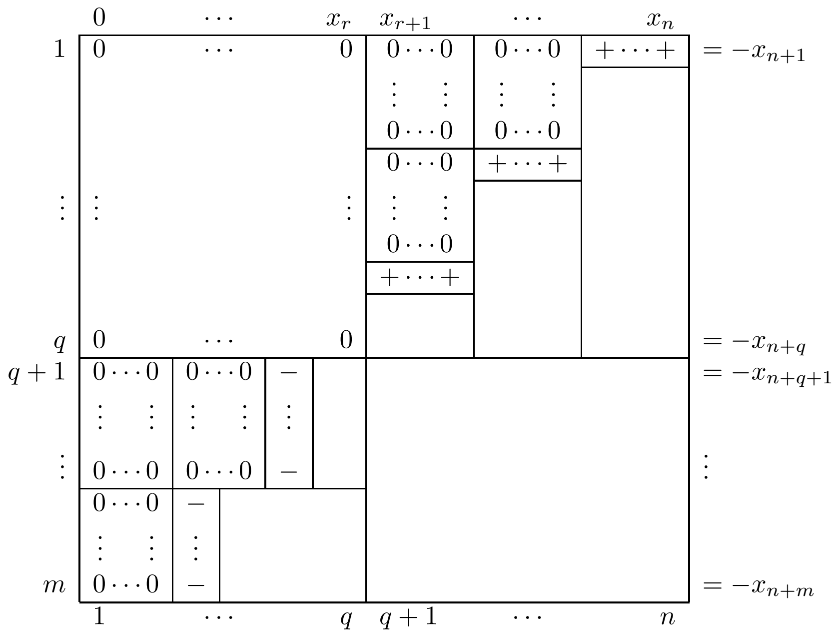

Figure 4.1: A Balinski-Tucker(B-T) Matrix

the Gaussian elimination method, this pivot operation can be described as follows. For a pivot on A_{pc}, the entry in position (p, c), augment the matrix A with the p-th unit vector (entry 1 in position p and 0's elsewhere). Now divide all entries of row p by A_{pc} and add multiples of this row to the entries of the other rows of A, in such a way that A_{pc} becomes the only nonzero entry in column c. Next, remove column c and replace it by the augmented column. This pivot operation is reversible: performing the pivot operation two times consecutively on the same entry restores the original matrix.

4.2.1 The connection between LO-problems and B-T matrix problems.

It is well known that the time complexity of solving an LO-problem belongs to the same complexity class as the problem of finding a feasible point for a linear constraint collection; see, for instance, Schrijver [47], chapter 10. In this section we will show that the B-T matrix problem has the same complexity as these two problems. The two problems can be formulated as follows.

The feasibility problem(FEAS):
Given a matrix $A \in \mathbb{R}^{m \times n}$ and a column vector $b \in \mathbb{R}^m$, test if $Ax \leq b$ has a solution, and if so, find one.

The LO-problem(LO):
Given a matrix $A \in \mathbb{R}^{m \times n}$, a column vector $b \in \mathbb{R}^m$, and a row vector $c \in \mathbb{R}^n$, test if $\max\{cx | Ax \leq b\}$ is infeasible, finite, or unbounded. If it is finite, find an optimal solution. If it is unbounded, find a feasible solution x_0, and a vector z with $Az \leq 0$ and $cz > 0$.

The proof that if one of these two problems can be solved in polynomial time, then also the other one can be found in Schrijver [47], chapter 10. In Schrijver's proof it is assumed that the entries in the matrices and vectors have rational values, because he uses the bit-size computational model. However, the proof is also valid for real matrices and vectors in the reals-computational model, and hence, the proof is also valid in case of strong polynomiality.

Theorem 4.1. *If one of the problems FEAS, LO, BT-MAT can be solved in (strongly) polynomial time, then so can each of the other problems.*

Proof. The equivalence of FEAS and LO is proved in Schrijver [47]. From Theorem 3.1 we know that if LO can be solved in (strongly) polynomial time, then also BT-MAT can be solved in (strongly) polynomial time, by taking a zero row for the objective row and a zero column for the rhs vector.

Next, we proof that if BT-MAT can be solved in (strongly) polynomial time, then FEAS can be solved in (strongly) polynomial time as well.

Assume that the B-T matrix problem can be solved in (strongly) polynomial time. Consider the following feasibility problem. For given A and b, find x such that $Ax \leq b, x \geq 0$ or prove that such an x does not exist. With the columns of A we associate the variables $x_1, \ldots, x_n$, and with the rows the 'slack-variables' $x_{n+1}, \ldots, x_{n+m}$. Let A' be the matrix A augmented with the rhs vector b with a negative sign; i.e. $A' = [A, -b]$.

We will treat $-b$ as a normal column of A' and associate with it the variable RHS. A' can be seen as the Tucker tableau of the inequalities $A'x \leq 0,\ x \geq 0$. Let $C = \{x \in \mathbb{R}^n \mid A'x \leq 0;\ x \geq 0\}$. Use a (strongly) polynomial algorithm to transform the matrix A' into an equivalent B-T matrix. Similar as with B-T tableaus, we can partition the variables into two sets. N is the set of variables that have zero values for all feasible points in C, and B is the set of variables for which all points in the relative interior of C have strictly positive values. Now, if RHS belongs to B, we can construct a feasible point v in the relative interior of the set described by $\{x \in \mathbb{R}^m, RHS \in \mathbb{R} \mid [A, -b][x, RHS]^T \leq 0, x \geq 0, RHS \geq 0\}$ in strongly polynomial time, in the same way as we did in the proof of Theorem 3.3.

Scaling v in such a way that the coefficient of RHS becomes one, yields a feasible point for the feasibility problem after removing the coefficient of RHS. If, on the other hand, RHS belong to N, then the feasibility problem FEAS has no solution, since any solution of the feasibility problem corresponds to a feasible point in C where the value of the RHS variable is fixed at the value one. $\square$

Theorem 4.1 shows that the B-T matrix problem belongs to the same time complexity class as the general LO-problem. The B-T matrix problem is a problem without a rhs vector, an objective vector, and without constraints. The matrix A can be seen as the coefficient matrix in a set of equations $x_D = -Ax_I$, where x_I is the vector of independent variables, and x_D the vector of dependent variables. In the description of BT-MAT, the operations EC, ER, and PIV only exchange rows and columns in a matrix. Therefore, the B-T matrix problem looks more like a sorting problem than an optimizing problem. Hence, the B-T matrix problem can be seen as a linear algebra or computer science problem, rather than as a problem in mathematical programming.

The B-T matrix problem is a special case of the following class of problems. If we have a real two-dimensional matrix A and we allow the three operations EC, ER, and PIV, as described in combination with the B-T matrix problem, that transform the matrix A into an equivalent matrix, what kind of matrix structures can be obtained?

From the theory of linear optimization, we know that it is possible to construct from A, by means of pivot operations, an equivalent matrix that has either a nonnegative row, or a nonpositive column, or both.

Equivalently, it is also possible to construct from A an equivalent matrix that has either a nonpositive row, or a nonnegative column, or both. The proof is left to the reader.

From our theory about B-T matrices we know that from A an equivalent B-T matrix can be constructed in polynomial time. Can this also be done in strongly polynomial time? If this would be possible, then LO-problems can be solved in strongly polynomial time as well. Therefore, we conjecture that if one wants to find a strongly polynomial algorithm for LO-problems, one should use the B-T matrix problem, because its description does not contain an objective function, a rhs vector, or even inequalities and nonnegativities. This 'restriction' does not make the problem easier, but no energy has to be spend on features that are not essential. Removing unnecessary details from a problem description, is one of the most important (and maybe difficult) tasks; see for instance Van Gasteren [27].

In the description of the B-T matrix problem, the operations ER and EC are very simple. The operation PIV is responsible for the difficulty of the problem. Although pivot operations are known for a few hundreds of years, the operation is not fully understood yet. In Nering & Tucker [45](page 138) one can read that: "It seems that there is something about pivoting that is only dimly understood at this point".

4.3 Interior Point Methods

Consider a primal LO-model in the following standard form

$$\min\{c^T x | Ax = b, x \geq 0\}, \tag{4.1}$$

where A is a $m \times n$ matrix with full row rank. In any interior point methods(IPM) the most dominant step is solving a linear equation system of the form

$$A(D^{(k)})^2 A^T u = v, \tag{4.2}$$

with v in $\mathbb{R}^m$ and u in $\mathbb{R}^n$; see for instance Roos, et al. [46]. $D^{(k)}$ is a diagonal scaling matrix with positive diagonal entries occurring at the k-th iteration of an IPM. The matrix $D^{(k)}$ depends on the IPM used. In case of primal IPMs one usually takes $D^{(k)} = X^{(k)}$, where $X^{(k)}$ is a diagonal matrix with diagonal vector $x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$, occurring at the k-the iteration of the algorithm. Let $N \subseteq \{1, 2, \ldots, n\}$ denote the set of indices of the coordinates of $x^{(k)}$ that are equal to zero on the optimal face. Similarly, let B denote the set of indices of the coordinates that have strictly positive values on the relative interior of the primal optimal face. Then (B, N) forms a partition of the column indices of A. This partition is called the optimal partition. Note that $x_N^{(k)} \rightarrow 0$ for any convergent IPM.

Convergence Assumption. We will assume that we are dealing with IPMs in which the matrix $D^{(k)}$ converges to a matrix D^* in such a way that the diagonal entries with indices in N converge to zero and the diagonal entries with indices in B converge to a strictly positive value.

If $D^{(k)}$ satisfies the Convergence Assumption, $A(D^{(k)})^2A^T$ converges to $[A_B(D_B^*)^2A_B^T \ , \ A_N(D_N^*)^2A_N^T] = [A_B(D_B^*)^2A_B^T \ , \ 0]$. Therefore, the limiting behavior of $A_B(D_B^{(k)})^2A_B^T$ determines the 'asymptotic' behavior of the equation system (4.2). Here, A_B and A_N denote the submatrices of A consisting of the columns of A with indices in B and N, respectively, and $D_B^{(k)}$ and $D_N^{(k)}$ denote the diagonal submatrices of $D^{(k)}$ consisting of the rows and columns with indices in B and N, respectively. With the usual definition of degeneracy nothing can be said about the behavior of the equation system (4.2) in the neighborhood of the optimum in the primal optimal face. Güler et al.[35] and Gondzio & Terlaky[31] write:

> "If the primal LO-model is degenerate and the dual model is nondegenerate, the matrix A_B has less than m columns and so rank$(A_B) < m$ and $A_B(D_B^{(k)})^2A_B^T$ is singular."

The definition for degeneracy in Güler et al.[35] reads as follows:

> "An LO-model is called degenerate if there exists a primal feasible x with less than m positive coordinates."

In the case that x is a vertex this definition of degeneracy coincides with the usual definition of degenerate vertex. However, if x is a feasible point that is not a vertex this definition becomes quite peculiar. If x is a point of a k-dimensional face F of the feasible region, it must have at least $n+1$ binding constraints. But this would mean that the degeneracy degree of F is at least $bnd(F, P) + dim(F) - n = n + 1 + k - n = k + 1$.

It seems that the authors of [35] did not realize this, since they never discussed degenerate points that are not vertices.

If the feasible region of the primal LO-model contains degenerate points, but none of them is located in or near the optimal face, then these points have no influence on the limiting behavior of the IPM. In this case A_B has at least m columns and $\mathrm{rank}(A_B) = m$. If the optimal face is degenerate according to our definition, the feasible region has degenerate points according to the definition in [35], but nothing can be said about the number of columns that form A_B. A_B may have much more than m columns and still have a rank smaller than m. As an example consider the following LO-model: min $\{x_5 \mid -x_5 + x_6 = 0;\ x_1+x_2+x_3+x_4+x_5+x_6 = 1;\ x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0\}$. Written in the form of an optimal B-T tableau, this model reads:

1	x_1	x_2	x_3	x_4	x_5	
0	0	0	0	0	1	$= -f$
0	0	0	0	0	-1	$= -x_6$
-1	1	1	1	1	-1	$= -x_7$

From this B-T Tableau the following facts can be concluded. The feasible point with $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0,\ x_7 = 1$ is optimal and degenerate with degeneracy degree one. Hence, "the model is primal degenerate". The dimension of the primal optimal face is four and the degeneracy degree is one. The optimal partition consists of $B = \{1, 2, 3, 4, 7\}$ and $N = \{5, 6\}$. The matrix A_B has five columns, one row, and its rank is equal to one.

With our definition of degeneracy, we are able to give a more precise characterization of the behavior of the rank of the equation system (4.2) in the neighborhood of the optimum in the primal optimal face.

Theorem 4.2. *Let the optimal face of the primal LO-model (4.1) have degeneracy degree σ, and let the matrix A have full row rank m. If $D^{(k)}$ satisfies the Convergence Assumption in the optimum, then the linear equation system (4.2) converges with increasing values of k to a system of which the rank of the coefficient matrix equals $m - \sigma$.*

Proof. We may assume that the tableau corresponding to LO-model (4.2) is already in the form of an optimal Balinski-Tucker(B-T) tableau.

A then has the following form:

$$A = \begin{array}{c} \begin{array}{cccc} B_1 & N_1 & N_2 & B_2 \end{array} \\ \left[\begin{array}{cccc} 0 & A_2 & I_\sigma & 0 \\ A_1 & A_3 & 0 & I_{m-\sigma} \end{array} \right] \end{array}.$$

The optimal partition (B, N) can easily be derived from the given optimal B-T tableau. In the matrix A, B_1 and B_2 are the subsets of B corresponding to the nonbasic and basic variables, respectively, and N_1 and N_2 are the subsets of N corresponding to the nonbasic and basic variables, respectively. Clearly, $|N_2|$ is equal to the degeneracy degree σ of the primal optimal face, and $|B_1|$ is equal to the dimension of the primal optimal face. Since $D^{(k)}$ converges to D^* and satisfies the Convergence Assumption, $AD^{(k)}$ will converge to

$$AD^* = \begin{array}{c} \begin{array}{cccc} B_1 & N_1 & N_2 & B_2 \end{array} \\ \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ A_1 D^*_{B_1} & 0 & 0 & I_{m-\sigma} D^*_{B_2} \end{array} \right] \end{array}.$$

with $D^*_{B_1}$ and $D^*_{B_2}$ being the nonsingular diagonal submatrices of D^* whose rows and columns correspond to B_1 and B_2 respectively. Hence, $I_{m-\sigma} D^*_{B_2}$ is a nonsingular diagonal matrix. Therefore, $\text{rank}(AD^*) = \text{rank}(I_{m-\sigma} D^*_{B_2}) = m - \sigma$. The rank of the matrix $A(D^{(k)})^2 A^T$ converges to the rank of $(AD^*)(AD^*)^T$ which is equal to $m - \sigma$. $\square$

In case of a degenerate optimal face, we have seen that the equation system (4.2) converges with increasing k to a singular system. So, one would expect numerical problems in the neighborhood of the optimal solution. However, if the IPM is implemented in a careful way, numerical instability does not occur. The reason for this can be found in Stewart[49] where it is proved that the norm of $(AD^2 A^T)^{-1} AD^2$ is bounded uniformly, independent of the scaling matrix D.

4.4 Degeneracy Graphs

Degeneracy in linear programming is a widely studied phenomenon, mostly because of its nasty aspects. It may cause, for instance, cycling in Simplex algorithms, and it may increase the complexity of applying sensitivity analysis. In Gal[18] degeneracy graphs were presented for the first time as a tool to study the properties of degenerate vertices. Since that time many researchers from the Fernuniversität Hagen in Germany study these graphs; see e.g. [18, 19, 20, 21, 22, 23, 41, 56, 25].

Degeneracy graphs have the feasible bases of a given LO-model as nodes, and the Simplex pivots as edges. The successive Simplex tableaus that are created during the execution of the Simplex algorithm form a path in such a graph. In this section we will study the relationships with B-T tableaus and solve a number of open problems from Kruse[41].

4.4.1 Representation and Degeneracy Graphs

Throughout this section we use the polyhedron representation P in the canonical form.

$$P = \{Ax \leq b; x \geq 0\}, x \in R^n, A \in R^{m \times n}, b \in R^m,$$

and we assume that $pol(P)$ is not empty. A basis B of P is denoted by a subset of $\{x_1, \ldots, x_{m+n}\}$, with $x_1, \ldots, x_n$ the entries of x, and $x_{n+1}, \ldots, x_{m+n}$ the slack variables of $Ax \leq b$.

Definition 4.1. *A representation graph is a graph G, of which the nodes are the feasible bases of P, and two nodes are connected with an edge iff the two corresponding bases can be obtained from each other by one pivot operation.*

Every feasible basis of P corresponds to a vertex of $pol(P)$, but a vertex of $pol(P)$ may correspond to several feasible bases, and hence to several nodes of the representation graph. The endpoints of an edge in G correspond to bases, that can be obtained from each other by one pivot step; an edge of the representation graph G is called a positive (negative) edge if the corresponding pivot is performed on a positive (negative) entry in the corresponding Simplex tableau. The subgraph of G that contains only the positive (negative) edges of G is called the positive (negative) representation graph, and is denoted by G_+ (G_-). In Theorem 4.4 it will be shown that G is always connected. However, G_+ and G_- may be disconnected.

Negative representation graphs

In general, if $pol(P)$ contains at least two vertices, G_- is disconnected, since it is not possible to go from one feasible basis in one vertex to another feasible basis in another vertex by negative pivots alone. Therefore, if G_- is connected, $pol(P)$ contains only one vertex. However, this is not sufficient for the connectness of G_-, as the following example shows.

Consider the polyhedron representation:

$$P = \{2x_1 \leq 0; x1 >= 0\}.$$

and use x_2 as the slack variable for the first constraint. The feasible tableaus are:

$$\begin{array}{r|c|c|} & rhs & x_1 \\ \hline g = & 0 & 0 \\ \hline -x_2 = & 0 & 2 \\ \hline \multicolumn{3}{c}{\text{Basis } B_1 : \{x_2\}} \end{array} \qquad \begin{array}{r|c|c|} & rhs & x_2 \\ \hline g = & 0 & 0 \\ \hline -x_1 = & 0 & 0.5 \\ \hline \multicolumn{3}{c}{\text{Basis } B_2 : \{x_1\}} \end{array}$$

The feasible region has only one vertex ($x_1 = x_2 = 0$) and two bases B_1 and B_2 that are connected in the representation graph G with an edge that corresponds to a positive pivot. The negative representation graph G_- consists of the two vertices of B_1 and B_2 and contains no edges. Therefore, G_- is not connected.

Positive representation graphs

In general, a positive representation graph will be connected. If no vertex of $pol(P)$ is degenerate, every feasible basis can be reached from any other feasible basis by means of positive pivots. This is the usual case in the Simplex method. It is however possible that the positive degeneration graph G_+ is not connected; see example (4.3). Below we give a number of properties of a constraint collection with a disconnected positive representation graph.

Definition 4.2. *Let v be any vertex of $pol(P)$. The degeneracy graph of v denoted by G^v, is the subgraph of G induced by the feasible bases of v. The subgraph of G^v that contains only positive (negative) edges is denoted by G^v_+ (G^v_-), respectively.*

Theorem 4.3. *Let G_+ be a disconnected positive representation graph of the constraint collection P. Then the following assertions hold.*
(a) For each vertex v of $pol(P)$, it holds that each connected component of G_+ contains a node corresponding to v.
(b) For each vertex v of $pol(P)$, G_+^v is disconnected, and every vertex of $pol(P)$ is degenerate.
(c) $pol(P)$ is unbounded.

Proof.
(a) Select an arbitrary connected component of G_+, and a node in it. Construct the Simplex tableau that corresponds to the selected node. Next, select an arbitrary vertex of $pol(P)$, and an objective function that is optimal in this vertex. Use a Simplex pivot rule, which uses only positive pivots to solve this tableau. The Simplex method creates a path in the selected component of G_+ that connects the selected node with a basis of the selected vertex. So, the arbitrary selected vertex has a basis in this component of G_+.
(b) Take any vertex v of $pol(P)$. Since G_+ is disconnected, and every component of G_+ contains a node that corresponds to a basis for v, G_+^v is also disconnected. Therefore, G_+^v must contain at least two bases. Hence v is degenerate.
(c) Take any vertex v of $pol(P)$, and let B_1 and B_2 be two bases of v located in different components of G_+. The basic variables that have a strictly positive value in v are in both bases, so that the difference between the two bases concerns only the basic variables with zero values. Let $Q = B_1 \setminus B_2$ denote the variables in B_1, but not in B_2, and let $R = B_2 \setminus B_1$. Suppose that the differences between B_1 and B_2 are as small as possible; i.e. with a positive pivot it is not possible to make Q and R smaller. We will show that the tableau corresponding to B_1 contains a column without positive entries, but with at least one negative entry, which means that P is unbounded. Let A_1 be the submatrix in the tableau of B_1 with columns corresponding to the nonbasic variables in R, and with rows corresponding to the basic variables in Q. Since B_1 and B_2 are both feasible bases, A_1 is a nonsingular submatrix, and because Q and R are as small as possible, A_1 has no positive entries, and every column of A_1 contains at least one nonzero entry. Let x_r be a nonbasic variable w.r.t. B_1 in R, and let x_q be a basic variable in Q that corresponds to a negative entry in the column of x_r in A_1.

Pivoting on this negative entry results in the basis $(B_1 \setminus \{x_q\}) \cup \{x_r\}$, which is a feasible basis, since the value of x_q is zero. If the column of x_r does not contain a positive entry, then the value of x_r can be increased unlimitedly, and therefore $pol(P)$ is unbounded. If, on the other hand, the column of x_r contains a positive entry, we can find a basic variable x_z, by means of the usual ratio test, such that a pivot on this positive entry yields the feasible basis $(B_1 \setminus \{x_z\}) \cup \{x_r\}$. After this pivot, the entry on the intersection of the column of x_z and the row of x_q has become positive. Pivoting on this positive entry results in the basis $(((B_1 \setminus \{x_z\}) \cup \{x_r\}) \setminus \{x_r\}) \cup \{x_z\} = (B_1 \setminus \{x_q\}) \cup \{x_r\}$, which is again a feasible basis. Since the difference between this basis and B_2 is smaller than the difference between B_1 and B_2, we have a contradiction with the assumption that the difference between B_1 and B_2 is as small as possible. Therefore, the column of x_r cannot contain a positive entry, and hence $pol(P)$ is unbounded. □

We will illustrate Theorem 4.3 by means of a small example. Consider the following polyhedron representation:

$$P = \{-x_1 \leq 0;\ -2x_1 + x_2 + x_3 \leq 1;\ x_1, x_2, x_3 \geq 0\}. \tag{4.3}$$

The feasible tableaus are (using x_4 and x_5 as slack variables):

	rhs	x_1	x_2	x_3
$g =$	0	0	0	0
$-x_4 =$	0	-1	0	0
$-x_5 =$	-1	-2	1	1

Basis $B_1 : \{x_4, x_5\}$

	rhs	x_1	x_2	x_5
$g =$	0	0	0	0
$-x_4 =$	0	-1	0	0
$-x_3 =$	-1	-2	1	1

Basis $B_2 : \{x_4, x_3\}$

	rhs	x_1	x_5	x_3
$g =$	0	0	0	0
$-x_4 =$	0	-1	0	0
$-x_2 =$	-1	-2	1	1

Basis $B_3 : \{x_4, x_2\}$

	rhs	x_4	x_2	x_3
$g =$	0	0	0	0
$-x_1 =$	0	-1	0	0
$-x_5 =$	-1	-2	1	1

Basis $B_4 : \{x_1, x_5\}$

	rhs	x_4	x_2	x_5
$g =$	0	0	0	0
$-x_1 =$	0	-1	0	0
$-x_3 =$	-1	-2	1	1

Basis $B_5 : \{x_1, x_3\}$

	rhs	x_4	x_5	x_3
$g =$	0	0	0	0
$-x_1 =$	0	-1	0	0
$-x_2 =$	-1	-2	1	1

Basis $B_6 : \{x_1, x_2\}$

The representation graph of this example is depicted in Figure 4.2.

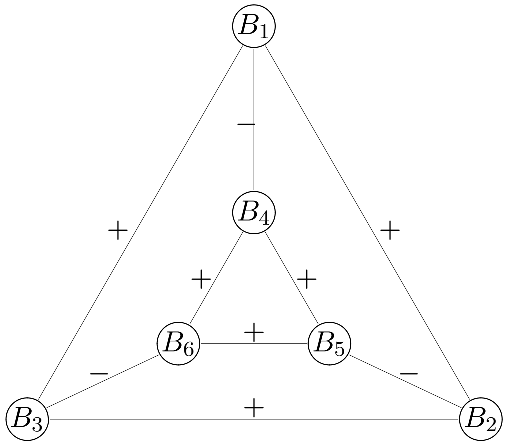

Figure 4.2: Representation graph of Example 4.3.

The vertices of the feasible region are $(x_1 = x_2 = x_3 = x_4 = 0, x_5 = 1)$, $(x_1 = x_2 = x_4 = x_5 = 0, x_3 = 1)$ and $(x_1 = x_3 = x_4 = x_5 = 0, x_2 = 1)$. It is easy to see that every vertex contains two bases and is degenerate. Namely, B_1 and B_4 correspond to $(x_1 = x_2 = x_3 = x_4 = 0, x_5 = 1)$, B_2 and B_5 correspond to $(x_1 = x_2 = x_4 = x_5 = 0, x_3 = 1)$, and B_3 and B_6 correspond to $(x_1 = x_3 = x_4 = x_5 = 0, x_2 = 1)$. Furthermore, the feasible region is unbounded because the column of x_1 in the tableau of B_1 contains no positive entry. Note that the positive representation graph consists of two disconnected triangles (B_1, B_2, B_3) and (B_4, B_5, B_6).

Theorem 4.4. *The representation graph G of a constraint collection P is connected.*

Proof. In the trivial case that G contains only one feasible basis, the theorem is obviously true. Assume that there are two arbitrary feasible bases, say B_1 and B_2. Start with a Simplex tableau with B_1 as basis and an objective function that is optimal for B_2. Use a Simplex pivot selecting rule, which uses only positive pivots to solve this tableau. The

Simplex method creates a path in G_+ that connects B_1 with a basis B_3, such that B_2 and B_3 correspond to the same vertex. Similar as in the proof of Theorem 4.3 a matrix A_1 can be constructed in the current tableau of B_3, being the intersection of the rows and columns that constitute the difference between B_2 and B_3. Performing pivot operations on positive and/or negative entries inside A_1 makes the difference between B_2 and the current B_3 smaller, until the current B_3 equals B_2. In this way a path from B_1 to B_2 is constructed. □

4.4.2 Optimal Degeneracy Graphs

Consider the LO-model

$$\max\{c^T x \mid Ax \leq b; x \geq 0\}, A \in R^{m\times n}, b \in R^m, c, x \in R^n. \tag{4.4}$$

Definition 4.3. *Let v be an optimal vertex of (4.4), and let B^v be the set of all optimal bases of v. The optimal-degeneracy(opt-deg) graph O^v is the subgraph of G^v induced by B^v. The subgraph of G_+^v (resp. G_-^v) induced by B^v that contains only positive (resp. negative) edges is called the positive (resp. negative) opt-deg graph of v and is denoted by O_+^v (resp. O_-^v).*

A natural question is whether it may happen that $G^v = O^v$. In Kruse[41], it is shown that it is possible that all bases of a degenerate vertex v are optimal. In his proof, Kruse gives an example in which the optimal vertex v has a degeneracy degree of one, and reports that examples with a degeneracy degree larger than one are not known. Kruse asks the question whether the assertion is true in general, or only under the condition that the degeneracy degree is one.

Kruse's question can be answered as follows. Actually, it is quite easy to construct examples with an arbitrary positive degeneracy degree. Namely, consider the following LO-model:

$$\begin{array}{llll}
\min & & y & \\
s.t. & Ax & & \leq 0 \\
 & & y & \leq 1 \\
 & x & & \geq 0 \\
 & & y & \geq 0
\end{array}$$

with $A \in R^{m\times n}$, $x \in R^m$, and $y \in R$. Obviously, the vertex with $y = 0, x = 0$ is optimal. Pivoting in A does not change the feasibility nor the optimality. A pivot in the column of y destroys the optimality. Therefore, all bases in the vertex with $x = 0, y = 0$ are optimal. The degeneracy degree of this vertex is m.

As a result of the duality between optimality and feasibility, any opt-deg graph is isomorphic with a representation graph, as the following theorem shows.

Theorem 4.5. *For any opt-deg graph O^v there is a constraint collection of which the representation graph, say G, is isomorphic with O^v; moreover, O^v_+ is isomorphic with G_-, and O^v_- is isomorphic with G_+.*

Proof. We start with an opt-deg graph in an optimal vertex v and construct a corresponding optimal Tucker tableau. Positive pivots in rows that correspond to basic variables with a positive optimal value are not allowed, since such pivots cause the leaving of the optimal vertex. Therefore these rows can be deleted from the Tucker tableau; we then obtain a tableau that corresponds to the same opt-deg graph. Call the underlying model of this tableau the primal model. This model contains the constraints that are binding in v, and is a representation of cone with v as top. The negative transpose of this tableau corresponds to a dual model for which the representation graph is isomorphic with the opt-deg graph of the primal model. Every pivot on a positive entry in the dual tableau transforms it into a tableau that is the negative transpose of the resulting primal tableau after a pivot on the corresponding positive entry in the primal tableau. Therefore, the positive opt-deg graph of the primal model is isomorphic with the negative representation graph of the dual model, and the negative opt-deg graph of the primal model is isomorphic with the positive representation graph of the dual model. □

Let P be a constraint collection and (LO) a corresponding LO-model. A constraint collection from Theorem 4.5 can be constructed as follows. Let P^v be the collection of constraints of P binding at an optimal vertex v. Note that these constraints form a cone with apex v. We dualize the modified (LO) with the constraint collection P^v instead of P. Then the dual constraint collection has a representation graph isomorphic with O^v.

We will illustrate the construction of such a constraint collection with the following example. The LO-model reads

$$\begin{array}{llllllll}
\max & & & & -4x_4 & & & \\
\text{s.t.} & & & & -x_4 & \leq & 0 & (\text{slack}: x_5) \\
& & -2x_2 & -x_3 & +4x_4 & \leq & 0 & (\text{slack}: x_6) \\
& -x_1 & -x_2 & +x_3 & +x_4 & \leq & 0 & (\text{slack}: x_7) \\
& & +x_2 & -3x_3 & & \leq & 2 & (\text{slack}: x_8) \\
& 3x_1 & +2x_2 & -5x_3 & -5x_4 & \leq & 3 & (\text{slack}: x_9) \\
& \multicolumn{7}{l}{x_1, x_2, x_3, x_4 \geq 0.}
\end{array}$$

The corresponding optimal Tucker tableau is

	1	x_1	x_2	x_3	x_4
$-f =$	0	0	0	0	4
$-x_5 =$	0	0	0	0	−1
$-x_6 =$	0	0	−2	−1	4
$-x_7 =$	0	−1	−1	1	1
$-x_8 =$	−2	0	1	−3	0
$-x_9 =$	−3	3	2	−5	−5

In the optimal vertex corresponding to this tableau, only the basic variables x_8 and x_9 have a strict positive value. After removing the rows of x_8 and x_9, and taking the negative transpose of the new problem, we obtain the following optimal dual tableau.

	1	y_5	y_6	y_7
$-g =$	0	0	0	0
$-y_1 =$	0	0	0	1
$-y_2 =$	0	0	2	1
$-y_3 =$	0	0	1	−1
$-y_4 =$	−4	1	−4	−1

The corresponding constraint collection of the dual model is

$$\begin{array}{llllll}
& & y_7 & \leq & 0 & (\text{slack}: y_1) \\
& 2y_6 & +y_7 & \leq & 0 & (\text{slack}: y_2) \\
& y_6 & -y_7 & \leq & 0 & (\text{slack}: y_3) \\
y_5 & -4y_6 & -y_7 & \leq & 4 & (\text{slack}: y_4) \\
\multicolumn{6}{l}{y_5, y_6, y_7 \geq 0,}
\end{array}$$

which has a representation graph that is isomorphic with the opt-deg graph of the LO-model we started with.

Theorem 4.5 immediately implies the following. First, opt-deg graphs are connected since representation graphs are connected. A proof of this fact is also given in Zörnig and Gal[56]. Secondly, positive opt-deg graphs are in general disconnected, except if all bases belong to one dual vertex, since negative representation graphs are in general disconnected, except if all bases belong to the same vertex.

A further question asked in Kruse[41] is whether lack of full dimensionality is a necessary condition for the disconnectness of negative opt-deg graphs. This question is answered by the following theorem.

Theorem 4.6. *If a negative opt-deg graph O_-^v of an LO-model is disconnected, then the feasible region of it is not fully dimensional.*

Proof. According to Theorem 4.5 a negative opt-deg graph is isomorphic with a positive representation graph. Theorem 4.3 shows that if a positive representation graph is disconnected, its corresponding polyhedron is unbounded. For this polyhedron there is a Tucker tableau that has a column without any positive entry, and at least one negative entry. Taking again the negative transpose of this tableau results in a tableau of the initial model that has a row with nonnegative entries. This means that the values of some variables are zero for every feasible point, and hence, the feasible region is not fully dimensional. □

The degeneracy graphs, as introduced in Definition 4.2, can be seen as special cases of opt-deg graphs, in which all bases in the specified optimal vertex are optimal. This can be accomplished by taking an all-zero objective function in a degenerate vertex. Therefore, the following theorem holds.

Theorem 4.7. *Let v be a vertex of the feasible region of an LO-model. For the degeneracy graph G^v, there is a polyhedron, represented by a constraint collection, that has a representation graph H isomorphic with G^v; moreover, G_+^v is isomorphic with H_-, and G_-^v is isomorphic with H_+.*

Proof. The proof is similar to the proof of Theorem 4.5. We use an all-zero objective function. Therefore, the objective row in an optimal tableau will have zero entries, as well as the right hand side in the negative transpose. □

So far we constructed representation graphs by means of a constraint collection. In general, we call a graph a representation graph if there exists a constraint collection that has a representation graph that is isomorphic with it. Similar definitions can be made for degeneracy graphs and optimal degeneracy graphs.

Corollary 4.1. *The classes of representation graphs, degeneracy graphs, and optimal degeneracy graphs are isomorphic.*

Proof. This result follows immediate from Theorem 4.5 and Theorem 4.7. □

Since the classes of representation graphs, degeneracy graphs, and optimal-degeneracy graphs are isomorphic, all properties of representation graphs can easily be translated to properties of degeneracy graphs and of opt-deg graphs. In order to find such properties for degeneracy graphs an opt-deg graphs, it suffices to construct the 'negative transposes' of properties of representation graphs.